IDD 6888
4/6/12
5B-C

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 10
1200 Sixth Avenue, Suite 900
Seattle, WA 98101-3140

OFFICE OF
COMPLIANCE AND ENFORCEMENT

Reply to: OCE-127

APR - 6 2012

Certified Mail Number 7011 1150 0000 7954 2421
Return Receipt Requested

FILE COPY

James Cagle, Risk Manager - EHS
Nu-West Industries, Inc.
Agrium Conda Phosphate Operations
3010 Conda Road
Soda Springs, Idaho 83276

Re: Background Soil Conditions and Proposed Soil Screening Levels, Nu-West Industries, Inc. Conda Phosphate Operations Facility, February 2, 2012; EPA Docket No. RCRA-10-2009-0186

Dear Mr. Cagle:

This letter is in response to the Background Soil Conditions and Proposed Soil Screening Levels report identified above (Report) that was submitted pursuant to the June 2009 Administrative Order on Consent (Order) issued under Section 3013 of the Resource Conservation and Recovery Act, Docket No. RCRA-10-2009-0186, and the additional radiological data provided by email to EPA on March 18, 2012.

Upon a thorough statistical review of the data provided, EPA has determined that the data sets submitted to our office reporting the analytical results from sampling of the three off-site background areas were of sufficient quality, such that the data sets in their entirety may be used for establishment of background concentration levels. The attached documents describe in greater detail EPA's statistical rationale, approach in the use of comparative values, and the screening value basis for each constituent of interest (COI). The approved screening values to be used are shown in the attached document, Nu-West Summary Screening Table, with the screening values and basis for each highlighted.

Nu-West's February 2, 2012 Report proposed a number of recommendations and approaches, which are addressed below.

Nu-West's Report states the following on page 6:

> Based on the data analysis presented above, it is believed the background data are appropriate for use. Pursuant to the Work Plan Addendum and as detailed below, WSP calculated the 95% upper confidence levels (UCLs) of the mean concentrations, identified 95% UCLs for evaluating potential impact, and compared available soil screening levels and the background concentrations to identify CVs.

EPA comment 1: EPA reviewed the sampling data and after performing statistical analyses, agrees that the sampling data collected are appropriate for use. However, use EPA's view is that the 95%-95% Upper Tolerance Limit (UTL 95-95) is the more appropriate statistic for this investigation. The attached document, Statistical Limits used to Estimate Background Threshold Value (BTV), explains in greater detail why UTL 95-95 was selected.

Nu-West's report states the following on page 7:

> For each constituent, WSP proposes to use the highest 95% UCL (Table 6) for the various depth intervals because this accounts for natural variations in the mineralogical content of soils in these areas, differences in depositional layers, and anthropogenic activities (e.g., automobile exhaust, addition of fertilizers, and mixing of the soil during tilling. Additionally, because the background concentrations established using IS are really central tendency estimates, using the highest of the 95% UCL estimates from the different soil intervals will somewhat reduce the chances of the decision error of identifying samples with concentrations in the range of background as being impacted.
>
> WSP also proposes that in instances where an offsite DU concentration exceeds the 95% UCL, further analysis be conducted to ascertain whether there is a statistically significant difference between the offsite DU and background means.

EPA comment 2: EPA reviewed the background sampling data, and found only a few COIs where the data indicated concentration variability with depth. These were barium, calcium, cadmium, manganese, potassium, sodium, and pH. There are no toxicity values for calcium, potassium or sodium. In the event that elevated levels of calcium, potassium or sodium are discovered, that alone will not drive follow-up actions, such as a risk assessment. The screening level for barium is driven by ecological risk, of which the value is higher than the UTL 95-95 value. For the remaining two COIs, cadmium and manganese, the calculated UTL 95-95 value was found to encompass all but a few data outliers.

In the event that sampling results indicate the exceedance of a screening value for any of the COIs, the analytical protocols in the approved Work Plan are to be followed, which include the following requirements in Section 4.2:

> All of the samples collected from 0 to 2 in-bgs and 2 to 6 in-bgs will be analyzed on receipt; samples from the remaining intervals will be held.[5] Following calculation of concentrations in the DUs, based on sample and triplicate results (Section 5), a comparison with the screening levels will be performed. If the calculated concentration for the 2 to 6 in-bgs sample aliquot from any DU exceeds a human health or ecological screening level, the sample from the next deepest interval (0.5 to 1 ft-bgs) will be analyzed, and so on until the concentration is below the screening level.
>
> [5] In the event the surficial soil samples are largely comprised of limestone or other discernible fill material, the next deeper sample will also be initially submitted for laboratory analysis to address the potential that the surficial material result are "masking" residual impacts in the subsurface.

Nu-West's report states the following on page 8:

> First, the Work Plan Addendum included EPA's Regional Screening Levels (RSLs) for residential and industrial exposure scenarios. The property on which the releases occurred or potentially occurred was purchased by Nu-West in December 2011. Consequently, WSP proposes eliminating the residential RSLs from consideration in the development of the human health risk-based soil screening levels.

EPA comment 3: EPA evaluated the residential screening levels, the industrial screening levels, Idaho Department of Environmental Quality Initial Default Target Levels (IDTLs), ecological risk levels to wildlife, and the background levels of the COIs, along with the administrative Order requirement that the Work Plan determine the nature and extent of potential contamination at or from the Facility. In a majority

of the cases, background is the driving factor for the screening level for the COIs. In circumstances where the ecological risk to wildlife was more stringent than the residential screening level for a given COI, but still greater than the measured background values in the off-site areas, then that ecological risk value was selected. Selection of a screening level greater than the IDTLs would potentially result in a lack of data to inform all parties on the nature and extent of potential contamination at or from the Facility, which is a requirement under the Order. Screening levels are not action or cleanup levels. In the event that any COIs exceed screening levels, a risk assessment and other actions would necessarily be undertaken prior to any potential corrective action.

Nu-West's report states the following on page 8:

> For parameters without Eco SSLs (thallium, fluoride, and uranium), the Work Plan Addendum proposed the Oak Ridge National Laboratory (ORNL) Ecological Screening Benchmarks (Table 7b). The ORNL Benchmarks include "wildlife" values that represent dietary concentrations food or prey consumed by various wildlife receptors. Using these values for the purposes of soil screening would require bioaccumulation factors for the various dietary items consumed by each wildlife receptor. Rather than going through this complex exercise, WSP proposes to use the ecological soil benchmarks for wildlife provided in a document also prepared for ORNL: Preliminary Remedial Goals for Ecological Endpoints. These values are presented in Table 7c.

EPA comment 4: The values represented in Table 7c show the lowest thallium value of 2.1 mg/kg. No values are shown for either fluoride or uranium in Table 7c. The human health screening level for thallium is 0.078 mg/kg. However, the test method detection limit for thallium is greater than the actual human health screening level. The background sampling indicated results that were non-detect. The fluoride risk level of 7 mg/kg is driven by the IDEQ IDTL, and is nearly double that of the measured background. Uranium is driven by the radionuclide criteria, which is more stringent than that of elemental uranium value.

Nu-West's report states the following on page 8:

> The background IS study was performed to ensure that regional conditions are taken into account in this process. Specifically, the background concentrations are to be compared to the risk-based levels to establish the CVs. Similar to the use of background concentrations for evaluating impacted areas, WSP proposes to use the highest 95% UCLs for the individual intervals as this would account for natural variations in the mineralogical content of soils in these areas, differences in depositional layers, and anthropogenic activities. Tables 8a and 8b present the updated human health and ecological screening values, the maximum calculated 95% UCLs, and identify the CVs.

EPA comment 5: See EPA comments 1 and 2 above. The enclosed documents provide additional detail as to why the 95% UCL is not appropriate for establishment of background.

The Off-Site Soil Sampling Plan Sampling and Analysis Work Plan Addendum, dated September 19, 2011, provided a schedule as Figure 11, with the different projects tasks scheduled on a week by week basis. As this EPA letter approves background concentration levels so that the next phase of work may proceed, EPA requests that Nu-West provide within twenty (20) calendar days a timeline and schedule for the completion of the remaining work activities in the Work Plan accordance with paragraph 62 of the Order.

If you have any questions on the information provided, please feel free to call me at (206) 553-2964. EPA staff are available to discuss any of the technical details. Alternatively, you may reach me via email at: Magolske.Peter@epamail.epa.gov. Thank you for your attention to this important matter.

Sincerely,

Peter Magolske
Air / RCRA Compliance Unit

Enclosures

cc: Brian Monson, Idaho Department of Environmental Quality
P. Scott Burton, Esq. Hunton and Williams LLP

Nu-West Summary Screening Table

Parameters	mg/kg BTV UTL95-95	Human Health Screening Levels EPA Soil RSLs Residential	Industrial	IDEQ REM IDTL/ Critical Pathway (d)	EPA Eco-SSLs (b) Soil Plants	Soil Invertebrates	Wildlife Avian	Wildlife Mammalian	PRGs (c) Wildlife Mammalian	RBTC (d) Wildlife Mammalian	Min Risk Screen	Basis	BTV	Screening Value	Basis
Metals															
Aluminum	15041	7,700	99,000								7,700	HH	15,041	15,041	Background
Antimony	0.5	3.1	41	4.77		78		0.27			.27	mammal	.50	.50	Background
Arsenic	4.695	0.39	1.6	0.39	18		43	46			.39	HH	4.70	4.7	Background
Barium	170.2	1,500	19,000	896		330		2,000			330	soil invert	170	330	Risk
Beryllium	0.89	16	200	1.83		40		21			1.63	IDEQ	.89	1.6	Risk
Cadmium	0.869	7	80	1.35	32	140	0.77	0.36			.36	mammal	.87	.87	Background
Chromium (total)	18.61	12,000	150,000	2,130			26	34			26	avian	18.61	26	Risk
Fluoride (total)	3.95	310	4,100	7.36						149	7.36	IDEQ	3.95	7	Risk
Iron	14811	5,500	72,000	5.76							5.76	IDEQ	14,811.00	14,811	Background
Lead	13.59	40	80	49.6	120	1,700	11	56			11.0	avian	13.59	14	Background
Manganese	742	180	2,300	223	220	450	4,300	4,000			180	HH	742.00	742	Background
Selenium	1.042	39	510	2.03	0.52	4.1	1.2	0.63			.52	plants	1.04	1	Background
Thallium	ND	0.078	1	1.55					2.1		.08	HH	ND	2	Risk
Vanadium	22.68	39	520				7.8	280			7.80	Avian	22.68	23	Background

Rads	pCi/g	HH PRG	Screening Value	Basis
Gross Alpha	6.210		6.210	Background
Gross Beta	5.520		5.520	Background
U-234	1.034	4.950	4.950	Risk
U-235	0.083	0.206	0.206	Risk
U-238	1.060	0.777	1.080	Background
Th-230	1.335	3.800	3.800	Risk
Ra-226	1.958	0.000	1.958	Background
Ra-228	1.756	0.033	1.756	Background
Pb-210	1.411	0.66	1.411	Background
Po-210	1.154	54.6	54.600	Risk
K-40	19.940	0.138	19.940	Background

http://rais.ornl.gov/cgi-bin/prg/PRG_search?select=rad

c: http://www.esd.ornl.gov/programs/ecorisk/documents/tm162r2.pdf
d: Booz Allen Hamilton 2011 RepA4-2101-020_rev
Metals without toxicity values were omitted: Ca, K, Mg, Na
http://rais.ornl.gov/cgi-bin/prg/PRG_search

Statistical Limits used to Estimate Background Threshold Value (BTVs)

Establishing a Background Data Set

- BTVs are estimated based upon "established" background data sets.
- Established Background Data Set: a data set representing background conditions free of any outliers potentially representing locations impacted by the site and/or other activities.
- Outliers when present in a data set result in inflated values of various statistics of interest including: UCL, UPL, UTL, and USL. The use of inflated statistics as BTV estimates will result in a higher number of false negatives.

False negative error rate (declaring a location clean when in fact it is contaminated) is controlled by making sure that one is dealing with a defensible background data set free of outliers and impacted site locations.

How BTVs are Used in the Decision Making Process?

- Onsite observations are compared with BTVs.
- Typically BTVs are used as not-to-exceed values (bright line); an onsite value exceeding a BTV may require additional investigation.
- Since BTVs are used as not-to-exceed values, UCL95 of mean are not used to estimate BTVs. None of the background guidance documents suggest the use of UCL95 to estimate BTVs (e.g., EPA 2002 Background Guidance for CERCLA Sites, 2009 Unified Guidance Document).
- Onsite values less than BTVs potentially represent unimpacted locations and are considered coming from the same background population.
- Onsite values exceeding BTVs potentially represent observations not belonging to the background population requiring further investigation.
- It is well known that the use of UPL95 as an estimate of BTV tends to incorrectly classify observations coming from the background population (comparable to background) as coming from the impacted site locations. The use of a UPL95 tends to result in a high number of false positives, especially when there are multiple COCs and many onsite observations are comapred with UPL95.
- Similarly by definition, the use of a UTL95-95 will classify at least 5% of background observations as coming from the impacted site locations.

- An observation coming from the background population will be less than or equal to USL95 with confidence coefficient of 0.95. By definition, USL95 does not exclude any observations coming from the background population when computing USL95. In order to provide proper balance between false positive and false negative error rates (especially when dealing with multiple contaminants) it is suggested to use USL95 to estimate BTVs.

Brief Description of Statistical Limits used to Estimate BTVs

1. **95% Upper Percentile,** $x_{0.95}$ of the background data set reprsents an estimate of the 95th percentile of the background population. 95% of the background data set values are $\leq$ $x_{0.95}$. It is expected that 95% of values coming from the background population will be less than or equal to $x_{0.95}$.

2. **95% Upper Prediction Limit ($UPL_1$95) for a single future observation** is designed to compare a *single future* observation with $UPL_1$95. We are 95% sure that a *"single"* future value from the background population will be less than or equal to $UPL_1$95 with confidence coefficient (CC) of 0.95. If an onsite value, $x_{onsite} < UPL_1$95, it is interpreted that x_{onsite} (=future value) comes from the background population with CC of 0.95. A $UPL_1$95 is not meant to perform more than 1 future comparison.

3. **95% Upper Prediction Limit (UPL_k95) for k ($\geq$1) future observations** is designed to compare *k future* observations with UPL_k95. We are 95% sure that *"k"* future values from the background population will be less than or equal to UPL_k95 with CC of 0.95. A UPL_k95 is meant to perform *k* future comparisons.

4. **95%-95% Upper Tolerance Limit (UTL95-95)** represents that statistic such that 95% observations from the target population (e.g., background) will be less than or equal to UTL95-95 with CC of 0.95. UTL95-95 is designed to simultaneously provide coverage for 95% of all potential observations (current and future) from the background population with CC of 0.95. UTL95-95 can be used to perform many onsite comparisons.

5. **95% Upper Simultaneous Limit (USL95)** represents that statistic such that all observations from the "established" background data set will be less than or equal to USL95 with CC of 0.95. It is expected that observations coming from the background population will be less than or equal to USL95 with 95% confidence. A USL95 can be used to perform many onsite comparisons.

Choosing Confidence Coefficient (CC)

- Higher statistical limits are associated with higher levels of CC. For an example, a 95% UPL is higher than a 90% UPL.

- Higher values of CC (e.g., 99%) tend to decrease the power of a test resulting in higher number of false negatives- dismissing contamination when present.

  Therefore, CC should not be set higher than necessary.

- Smaller values of CC (e.g., 0.80) tend to result in higher number of false positives (e.g., declaring contamination when not present).

- In most practical applications, choice of 95% CC provides a good compromise between confidence and power.

Sample Size

- Smaller sample sizes (e.g., <10-15) tend to yield estimates with higher variabilities, which in turn result in higher values of USLs, UTLs and UPLs.

- Higher level of uncertainty in a background data set (e.g., due to a smaller background data set) tends to dismiss contamination as representing background conditions (results in more false negatives, i.e., identifying a location that may be dirty as background).

- Computation of nonparametric upper limits (e.g., UPLs, UTLs, and USLs) based upon order statistics require lager data sets. Nonparametric upper limits based upon order statistics (e.g., Maximum or second maximum) may not provide the desired coverage as they do not take data variability into account.

95%-95% Upper Tolerance Limit (UTL95-95)

A UTL *(1-α)-p* (e.g., UTL95-95) based upon an established background data set represents that limit such that *p%* of the sampled data will be less than or equal to UTL with CC, (1-α). It is expected that p% of the observations belonging to the background population will be less than or equal to UTL with CC, (1-α). A UTL *(1-α)-p* represents a *(1 – α) 100%* upper confidence limit for the p^{th} percentile of the underlying background population.

Interpreting UTL95-95

- UTL95-95 based upon a background data set represents that limit such that 95% observations from the background population will be less than or equal to that limit with confidence coefficient of 0.95. A UTL95-95 represents a 95% upper confidence limit for the 95% percentile.

- UTL95-95 is that value which will be exceeded by all values potentially coming from the background population less than 5% of the time with confidence coefficient 0.95.

- For a UTL95-95, 5 exceedances per 100 comparisons (of background values) can result just by chance for overall CC of 0.95; or 5% exceedances (in any number of comparisons) can occur just by chance with overall CC of 0.95. Similarly, for UTL95-99, 1 exceedance per 100 comparisons can result just by chance for overall CC of 0.95.

- Parametric UTLs take variability into account.
- When sample size is large (e.g., 500), UTL95-95 approaches the upper 95th percentile; UTL90-90 will approach the upper 90th percentile.

Use of UTL95-95 is preferred to $UPL_1 95$ when the number of future comparisons is large and/or unknown.

95% Percentiles to Estimate BTV

Based upon an established background (e.g., RBRA) data set, the 95th percentile, $x_{0.95}$ represents that statistic such that 95% of observations from the background data set are less than or equal to $x_{0.95}$. By definition, about 5% of values from the background data set will exceed $x_{0.95}$.

Interpreting 95% Percentile

- $x_{0.95}$ represents an estimate of the of the 95th percentile of the background population. It is expected that 95% of observations (values) coming from the background population will be less than or equal to $x_{0.95}$.
- If one uses 95% percentile to estimate BTV, at least 5% onsite observations with concentrations comparable to background will be determined as not belonging to the background population even when they actually come from the background population.
- If an onsite value, x_{onsite} exceeds background $x_{0.95}$, it may be concluded that x_{onsite} does not belong to the background population.
- 95% percentile, $x_{0.95}$ does not take variabilities of future observations into account. The use of 95% percentile to estimate BTV potentially may lead to a higher number of false positives resulting in unnecessary cleanup (i.e., determining a clean onsite location comaparable to background as dirty).

Upper Prediction Limits (UPLs) and Upper Tolerance Limits (UTLs) to Estimate BTV

- Unlike 95% percentile, UPLs and UTLs provide predictive setup for future observations.
- Parametric UPLs and UTLs account for extra amount of variability to accommodate future observations.

UPLs

UPLs can be computed for 1 or more future observations (e.g., future onsite values). Let UPL_k 95 represent a 95% UPL for *k (≥1)* future observations. A UPL_k 95 is designed to provide coverage for *k* future observations with confidence coefficient (CC) 0.95.

Interpreting $UPL_1$95 (comparing 1 single observation)

- We are 95% sure that a *"single"* future value from the background population will be less than $UPL_1$95 with CC= 0.95. If an onsite value, $x_{onsite} < UPL_1$95, it is interpreted that x_{onsite} (=future) value comes from the background population with CC of 0.95.
- An onsite value, x_{onsite} exceeding $UPL_1$95 potentially represents a value not belonging to the background population.
- UPLs are useful when background data set is of smaller size (e.g., <10-15); and/or a few and known number of future observations are to be compared with a UPL.

Improper use of $UPL_1$95 to perform many future comparisons

In practice users tend to use $UPL_1$95 for many future comparisons which results in a higher number of false positives (observation declared contaminated when in fact they are clean). When *k* future comparisons are made with a UPL_1, some of those future observations will exceed UPL_1 just by chance, each with probability 0.05. For proper comparison, UPLs need to be computed according to the number of comaprisons that will be performed.

In order to achieve the specified false rejection rate of 0.05, need to take the number of future comparisons into account.

If many (e.g., =30) independent onsite comparisons (e.g., Ra-226 activity from 30 onsite locations) are made with the same UPL_1, each onsite value may exceed that UPL_1 with probability 0.05 just by chance. The overall probability of at least one of those 30 comparisons being significant (exceeding BTV) just by chance is

$\alpha_{actual} = 1-(1-\alpha)^k = 1 - 0.95^{30} \sim 1-0.21 = 0.79$ (false positive rate).

This means that the probability (overall false positive rate) is 0.79 that at least one of the 30 onsite locations will be considered contaminated even when they are comparable to background.

Similar arguments hold when multiple (=*m*) contaminants are analyzed, and status (clean or impacted) of an onsite location is determined based upon *m* comparisons (one for each analyte).

UPL_k 95 for k future comparisons

In order to achieve the specified false positive error rate, α , UPL_k for *k* future observations should be used to estimate the BTV when *k* comparisons are to be performed with the BTV. A UPL_k uses appropriate critical value (based upon Bonferroni inequality) to accommodate k future observations. These UPLs satisfy the relationship: $UPL_1 \leq UPL_2 \leq UPL_3 \leq \ldots \leq UPL_k \ldots$

A UPL_{30} for 30 future observations is given by:

$$UPL_{30}95 = \left(\bar{x} + t_{((1-0.05/30),n-1)} s \sqrt{1+\frac{1}{n}} \right)$$

95% Upper Simultaneous Limit (USL95)

A *(1 - α)100%* USL based upon an "established" background data set is meant to provide coverage for all observations, x_i, $i = 1, 2, n$ simultaneously in the background data set. The two-sided *(1 - α)100%* simultaneous interval based upon the first order Bonferroni inequality (Singh and Nocerino, 1997) is given as follows:

$$P\left(\bar{x} - sd_\alpha^b \leq x_i \leq \bar{x} + sd_\alpha^b; i := 1,2,\ldots,n\right) = 1 - \alpha .$$

Here, $(d_\alpha^b)^2$ represents the critical value (obtained using the Bonferroni inequality) of the maximum Mahalanobis distance (Max (MDs)) for α level of significance.

Similarly, a one-sided *(1 - α)100%* USL given by:

$$P\left(x_i \leq \bar{x} + sd_{2\alpha}^b; i := 1,2,\ldots,n\right) = 1 - \alpha \ ;$$

Here $(d_{2\alpha}^b)^2$ is the critical value of Max (MDs) for $2^*\alpha$ level of significance.

Interpreting USL95

- A USL95 based upon an established background data set represents that estimate of BTV such that all observations in the background data set are less than or equal to USL95 with 95% CC.
- It is expected that observations coming from that background population will be less than or equal to USL95 with 95% CC.

- A USL can be used when many (and/or unknown) future onsite observations need to be compared with BTV.
- The false positive error rate does not change with the number of comparisons, as USL95 is meant to perform many comparisons simultaneously.
- USLs take variability into account.